

Exercise 8

Lecture 11: Neural Networks

Scalable Data Engineering

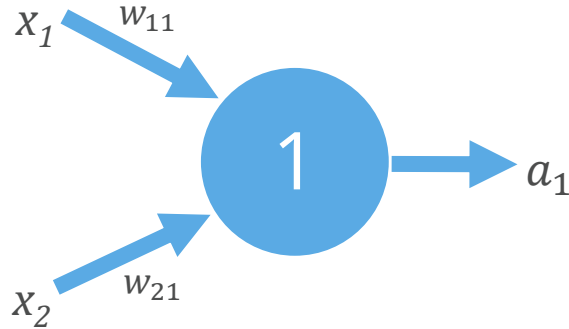
7.1 Are the following statements true or false? Give reasons for your answers.

- a) Supervised learning does not need labeled data.
- b) The output of a neuron is the weighted sum of all its inputs.
- c) A perceptron is a basic model for a neuron.
- d) To build a non-linear predictor, one can connect layers of perceptrons with linear activations to a multilayer perceptron.
- e) Stochastic gradient descent only uses one input sample for backpropagation gradient optimization.
- f) Backpropagation is based on four basic derivatives.
- g) Exploding gradients means large errors E lead to small to no changes in weights.
- h) The main problem of ReLU activations are „hyperactive“ neurons.
- i) Besides regularization, Dropout can be used to prevent overfitting in neural networks.

Perceptron Training

7.2 Train the OR-Perceptron until convergence (all $a_1 = t$).

- 1 is the step activation function
- $\alpha = 0.3$



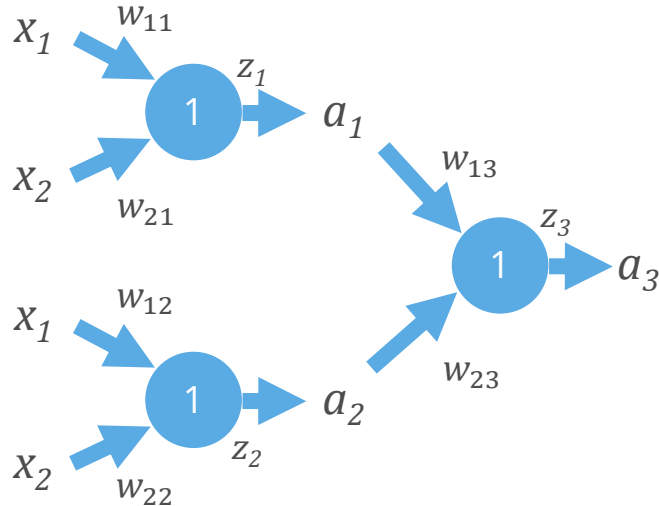
$$w_{ij}^{new} = w_{ij}^{old} + \alpha \cdot (t_j - a_j) \cdot x_i$$

x_1	x_2	t	z_1	a_1
1	0	1		
0	1	1		
0	0	0		
1	1	1		

Initial w_{ij}	w_{11}	w_{21}
	-0.2	0.4

7.3 Train the Multilayer Perceptron to produce the correct outputs.

- 1 is the step activation function
- $\alpha = 0.5$



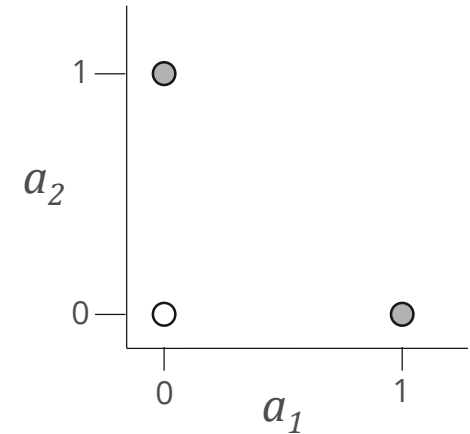
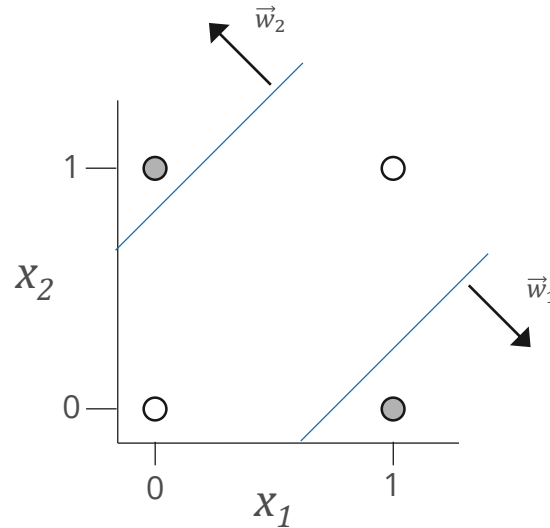
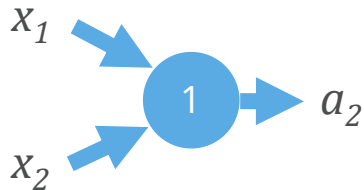
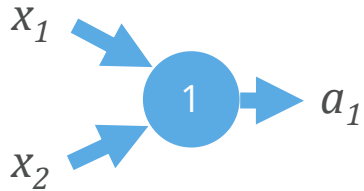
x_1	x_2	t_{a_1}	t_{a_2}	t_{a_3}	z_1	a_1	...
1	0	1	0	1			
0	1	0	1	1			
0	0	0	0	0			
1	1	0	0	0			

Initial w_{ij}	w_{11}	w_{21}	w_{12}	w_{22}	w_{13}	w_{23}
	0.5	-0.5	0.5	-0.5	0.5	0.5

Perceptron Training

Perceptron

- Linear separability
- How to model XOR?

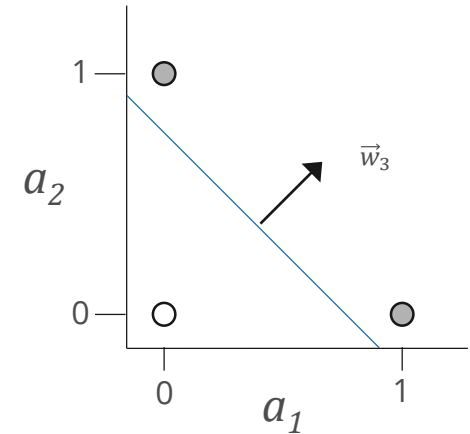
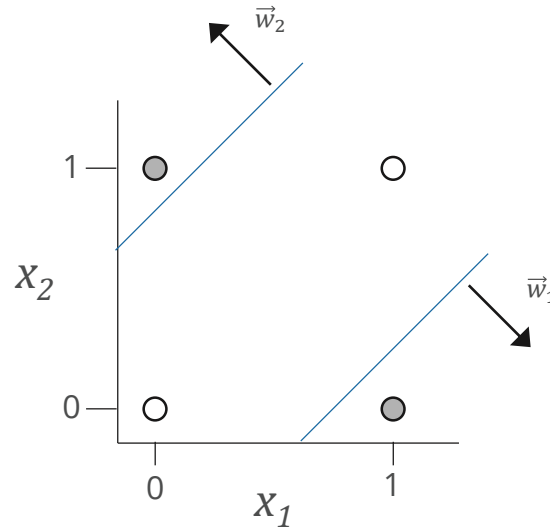
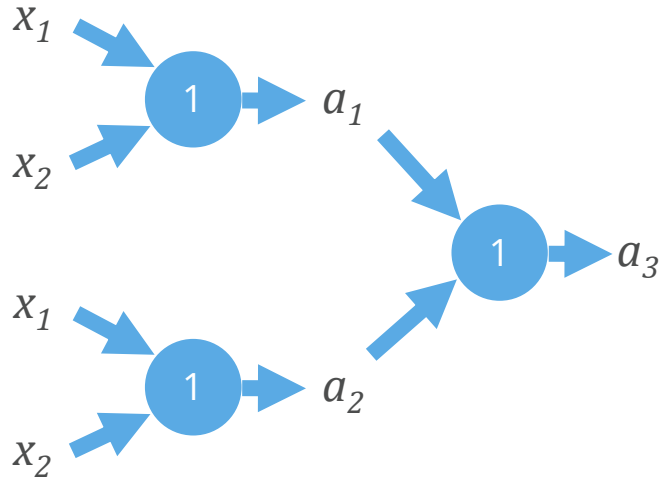


x_1	x_2	a_1	a_2
0	0	0	0
1	0	1	0
0	1	0	1
1	1	0	0

Perceptron Training

Perceptron

- Linear separability
- How to model XOR?



x_1	x_2	a_1	a_2	a_3
0	0	0	0	0
1	0	1	0	1
0	1	0	1	1
1	1	0	0	0